

1 Title Page

2 Title

3 Baseline expression of immune gene modules in blood is associated with primary response
4 to anti-TNF therapy in Crohn's disease patients

5 Authors

- 6 Benjamin Y. H. Bai (ORCID: 0000-0003-0765-6398) [1, 2]
7 Mark Reppell [3]
8 Nizar Smaoui [3]
9 Jeffery F. Waring [3]
10 Valerie Pivorunas [3]
11 Heath Guay [3]
12 Simeng Lin [4, 5]
13 Neil Chanchlani [4, 5]
14 Claire Bewshea [5]
15 James R. Goodhand [4, 5]
16 UK Inflammatory Bowel Disease Pharmacogenetics Study Group
17 Nicholas A. Kennedy [4, 5]
18 Tariq Ahmad [4, 5]
19 Carl A. Anderson [1]

20 **Affiliations**

- 21 1. Genomics of Inflammation and Immunity Group, Wellcome Sanger Institute,
22 Hinxton, UK;
- 23 2. Postgraduate School of Life Sciences, University of Cambridge, Cambridge, UK;
- 24 3. AbbVie Inc., Chicago, IL, USA;
- 25 4. Department of Gastroenterology, Royal Devon University Healthcare NHS
26 Foundation Trust, Exeter, UK;
- 27 5. Exeter Inflammatory Bowel Disease and Pharmacogenetics Research Group,
28 University of Exeter, Exeter, UK.

29 **Short title for running head**

30 Baseline module expression associated with anti-TNF response

31 **Abbreviations**

- 32 ADA: adalimumab
- 33 AIC: Akaike information criterion
- 34 BMI: body mass index
- 35 BTMs: blood transcriptomic modules
- 36 CD: Crohn's disease
- 37 CRP: C-reactive protein
- 38 FDR: false discovery rate
- 39 IBD: inflammatory bowel disease
- 40 IFX: infliximab

- 41 IMID: immune-mediated inflammatory disease
- 42 LOR: loss of response
- 43 PANTS: Personalised Anti-TNF Therapy in Crohn's Disease
- 44 PNR: primary non-response/non-responder
- 45 PR: primary response/responder
- 46 TNF: tumour necrosis factor
- 47 UC: ulcerative colitis
- 48 rRNA: ribosomal RNA
- 49 ssGSEA: single-sample gene set enrichment score

50 **Address for correspondence**

- 51 Carl A. Anderson
- 52 Genomics of Inflammation and Immunity Group, Wellcome Sanger Institute, Hinxton, UK
- 53 CB10 1SA
- 54 Tel: +44 (0)1223 834244
- 55 Fax: +44 (0)1223 494919
- 56 ca3@sanger.ac.uk

57

58 Abstract

59 **Background and Aims.** Anti-TNF therapy is widely used for treatment of inflammatory
60 bowel disease (IBD), yet many patients are primary non-responders (PNR), failing to respond
61 to induction therapy. We aimed to identify blood gene expression differences between
62 primary responders (PR) and primary non-responders to anti-TNF monoclonal antibodies
63 (infliximab and adalimumab); and to predict response status from blood gene expression
64 and clinical data.

65 **Methods.** The Personalised Anti-TNF Therapy in Crohn's Disease (PANTS) study is a UK-wide
66 prospective observational cohort study of anti-TNF therapy outcome in anti-TNF naive
67 Crohn's disease (CD) patients (ClinicalTrials.gov identifier: NCT03088449). Blood gene
68 expression in 324 unique patients was measured by RNA-seq at baseline (week 0), and at
69 weeks 14, 30, and 54 after treatment initiation (total sample size = 814).

70 **Results.** After adjusting for clinical covariates and estimated blood cell composition,
71 baseline expression of MHC, antigen presentation, myeloid cell enriched receptor, and other
72 innate immune gene modules was significantly higher in anti-TNF responders. Expression
73 changes from baseline to week 14 were generally of consistent direction across patients but
74 with greater magnitude in responders, with the exception of interferon-related genes that
75 were upregulated uniquely in non-responders. Expression differences between responders
76 and non-responders observed at week 14 were maintained at week 30 and week 54.
77 Prediction of response status from baseline clinical data, cell composition, and module
78 expression was poor.

79 **Conclusions.** Baseline gene module expression was associated with primary response to
80 anti-TNF therapy in PANTS patients. However, these expression differences were insufficient
81 for clinically useful prediction of response.

82 **Keywords**

83 Anti-TNF, transcriptomics, Crohn's disease

84

85 Introduction

86 Crohn's disease (CD) is a chronic immune-mediated inflammatory disease (IMID) of the
87 gastrointestinal tract. Along with ulcerative colitis (UC), it is one of the two main forms of
88 inflammatory bowel disease (IBD). The development of anti-TNF biologic therapies has
89 revolutionised patient care for CD and a number of other IMIDs over the last two decades.
90 Two major anti-TNF drugs, infliximab and adalimumab, are IgG1 monoclonal antibodies that
91 bind both soluble and transmembrane TNF, inhibiting their interactions with TNF
92 receptors.^{1,2} Two main mechanisms of action have been proposed: induction of CD4⁺ T cell
93 apoptosis in the gut mucosa by inhibiting the TNF-TNFR2 interaction; and binding of the
94 antibody tail (Fc region) of the drug to Fc receptors on monocytes, inducing their
95 differentiation into wound-healing M2 macrophages.^{3,4}

96
97 Unfortunately, anti-TNF therapy is not always effective at treating IBD. Various types of
98 treatment failure can occur: primary non-response (PNR) within the induction period (the
99 first 12–14 weeks for infliximab and adalimumab), secondary loss of response (LOR) during
100 maintenance therapy after an initial response, failure to achieve remission after the
101 treatment course, or adverse events that lead to treatment discontinuation.⁵ For IBD
102 patients, the incidence of PNR is 10–40%, and the incidence of secondary LOR among initial
103 responders is 24–46% in the first year of treatment.^{6–8} The ability to predict PNR and LOR
104 could help guide changes in treatment regimens, such as dose intensification or switching to
105 a drug class with a different mechanism of action.^{2,6} Reliable baseline prediction would be
106 especially valuable, allowing stratification of patients to effective therapies from treatment
107 initiation, minimising healthcare costs and patient burden.

108

109 Clinical variables reported to be associated with anti-TNF response include age, disease
110 duration, body mass index (BMI), smoking, C-reactive protein (CRP) levels, faecal
111 calprotectin levels, serum drug concentrations, and anti-drug antibody concentrations.
112 These associations have mostly been reported in small retrospective cohorts and have
113 rarely been independently validated.^{7,9-13} In the Personalised Anti-TNF Therapy in Crohn's
114 Disease (PANTS) study, the largest study of infliximab and adalimumab response in CD
115 patients to date (enrolment $n = 1610$), baseline obesity, smoking, and greater disease
116 activity were associated with low serum drug concentration after induction. Low drug
117 concentration was in turn associated with PNR and non-remission, suggesting
118 immunogenicity may be mediating treatment failure by increasing drug clearance rate.⁸

119

120 Multiple studies have attempted to define transcriptomic predictors for anti-TNF
121 response.¹²⁻²⁰ Reported associations include *TREM1* expression, identified as a marker of
122 anti-TNF response in different studies with inconsistent directions of effect. Gaujoux et al.¹⁶
123 found *TREM1* expression was lower in gut biopsies from infliximab responders than in non-
124 responders (total cohort size $n = 72$), but higher in responders in a separate cohort
125 measuring baseline whole blood expression ($n = 22$). By contrast, Verstockt et al.¹⁸ reported
126 lower *TREM1* expression in responders to infliximab and adalimumab in both baseline gut
127 biopsies ($n = 44$) and baseline whole blood ($n = 54$). Proposed reasons for the discrepancy
128 include false positives due to small sample sizes, differences in patient ethnicity, and
129 differing definitions of response.^{12,21} In general, small sample sizes, and variation among
130 studies in analysis methods, anti-TNF drug, response definition, tissues sampled, and
131 disease subtype make a consensus hard to establish. Few markers for anti-TNF response of

132 any type, clinical or transcriptomic, have been validated in independent studies, and none
133 have yet been translated to routine clinical practice.¹³

134

135 To identify novel transcriptomic associations with primary response to anti-TNF therapy, we
136 generated longitudinal RNA-seq data from peripheral blood samples taken from a subset of
137 the PANTS cohort (182 PR, 142 PNR) during the first year of follow-up. Differential gene
138 expression (DGE) between primary responders and non-responders was performed at
139 baseline (week 0), post-induction (week 14), and during maintenance (week 30 and week
140 54). We detected differences in gene module expression that may reflect differences in
141 disease characteristics or severity that influence risk of primary non-response. As this is one
142 of the largest datasets currently available for assessing transcriptomic associations with
143 anti-TNF response in IBD, we also examined the significance of previously reported
144 transcriptomic markers from the literature. Finally, we evaluated the utility of measuring
145 module expression for prediction of primary response status.

146 **Materials and Methods**

147 **Study design**

148 PANTS is a UK-wide, multicentre, prospective observational cohort study reporting the
149 treatment failure rates of the anti-TNF drugs infliximab (originator, Remicade [Merck Sharp
150 & Dohme, Hertfordshire, UK] and biosimilar, CT-P13 [Celltrion, Incheon, South Korea]), and
151 adalimumab (Humira [AbbVie, Cambridge, MA]) in 1610 anti-TNF-naive patients with active
152 luminal CD. The study design has been described in detail previously.^{8,55} In brief, patients
153 were recruited at the time of first anti-TNF exposure between February 2013 and June 2016,
154 and evaluated for 12 months or until drug withdrawal. Eligible patients were aged ≥ 6 years
155 with evidence of active luminal Crohn's disease involving the colon and/or small intestine.
156 Four major study visits were scheduled at first dose (baseline, week 0), post-induction (week
157 14), week 30, and week 54. Additional visits were scheduled at treatment failure or exit. At
158 baseline, clinical and demographic data were recorded, including sex, ethnicity, BMI,
159 smoking status, age at diagnosis, disease duration, Montreal classification, prior medical and
160 drug history, and previous Crohn's disease-related surgeries. At every visit, disease activity
161 score, weight, current therapy, and adverse events were recorded.

162 **RNA-seq sample selection**

163 A subset of PANTS patients was selected for RNA-seq, with the inclusion criteria: age ≥ 16
164 years; and baseline CRP ≥ 4 mg/L and/or baseline calprotectin > 100 $\mu\text{g/g}$. The target sample
165 size was 200 patients on infliximab and 200 patients on adalimumab, with an even split

166 between PR and PNR within each drug group. PR and PNR were defined based on Kennedy
167 et al.⁸ patient outcome criteria:

- 168 • *Primary non-response (assessed at week 14)*: exit before week 14 because of
169 treatment failure (including resectional inflammatory bowel disease surgery) or
170 corticosteroid use at week 14 (new prescriptions or if previous dose had not been
171 stopped). Patients whose CRP did not decrease to 3 mg/L or less or by 50% or more
172 from baseline (week 0), and whose HBI score did not decrease to 4 points or less or
173 by 3 points or more from baseline, were also classified as having a primary non-
174 response.
- 175 • *Primary response (assessed at week 14)*: decrease in CRP to 3 mg/L or less or by 50%
176 or more from baseline (week 0) and a decrease in HBI to 4 points or less or by 3
177 points or more from baseline.
- 178 • *Remission (assessed at weeks 14, 30, 54; implies primary response)*: CRP of 3 mg/L or
179 less and HBI score of 4 points or less, no ongoing steroid therapy, and no exit due to
180 treatment failure.

181 Steroid use was defined as any systemic therapy, either oral or intravenous (including use of
182 steroids for other conditions), but excluding single pre-infusion dosing with hydrocortisone.

183

184 PNR were required to exhibit primary non-response at week 14 and non-remission at week
185 54. PR were required to exhibit primary response or remission at week 14, and be in
186 remission at week 54 (or week 30 if week 54 status was unknown). Furthermore, within
187 infliximab-treated patients, PNR and non-PNR were matched based on baseline
188 immunomodulator use, baseline steroid use, age at first dose, baseline albumin, sex, and
189 weight at study entry.

190 **Whole blood RNA-seq**

191 Whole blood was collected in RNA Tempus tubes (Applied Biosystems) and stored at -80°C
192 until extraction (QIAsymphony PAXgene Blood RNA Kit, Qiagen). RNA was quantified using
193 the QuBit BR RNA (ThermoFisher), and RNA integrity was assessed with the 4200
194 TapeStation (Agilent). RNA-seq libraries were prepared using the Kapa mRNA HyperPrep Kit,
195 with depletion of ribosomal RNA (rRNA) and globin mRNA using the QIAseq FastSelect RNA
196 Removal Kit, and adapter ligation with IDT xGEN Dual Index UMI adapters. A total of 1 141
197 samples from 396 patients were sequenced. Raw sequencing data was demultiplexed with
198 Picard²² and aligned to the reference genome (GRCh38) using STAR (v2.6.1d)²³. Reads were
199 deduplicated using UMI-tools²⁴ and quantified against the Ensembl 96 gene annotation with
200 featureCounts (v1.6.4).²⁵

201

202 Outlier samples were excluded, defined as > 2 standard deviations from the mean based on
203 percentage of aligned reads in coding regions reported by Picard, percentage of unique
204 reads and number of unique reads. Samples with sex mismatch against documented sex
205 were removed. As gene expression measured from bulk tissue is heavily dependent on cell
206 composition,²⁶ cell proportions of six common cell types in whole blood (CD4⁺ T cells, CD8⁺ T
207 cells, B cells, NK cells, monocytes, and granulocytes) were estimated using the Houseman
208 method²⁷ from paired DNA methylation data.²⁸ Samples missing clinical data and/or cell
209 proportion estimates were removed. A total of 814 samples remained after filtering. To
210 accommodate variability in sampling day, samples were mapped to timepoints based on
211 Kennedy et al.⁸ windows around major visits: week 0 (week -4-0), week 14 (week 10-20),
212 week 30 (week 22-38), and week 54 (week 42-66). Samples taken at additional visits (LOR

213 or exit) falling within one of the windows were mapped to that timepoint, unless the patient
214 also had a major visit sample inside that window. The mapping of samples to timepoints is
215 shown in Fig. S1a. The number of samples per patient ranged from one to four, with a
216 median of three (Fig. S1b).

217

218 Counts were normalised for library size using edgeR (v3.28.1).²⁹ Globin genes and short non-
219 coding RNAs were removed, and genes with low expression were filtered, requiring genes to
220 have at least 1.25 counts per million in > 10% of samples and non-zero expression in > 90%
221 of samples. Expression data from 15 511 genes remained after filtering. Finally, log₂
222 expression values were generated using variancePartition/voom.^{30,31}

223 **Statistical analyses**

224 A full description of the statistical analyses can be found in the Supplementary Methods. In
225 brief, DGE analyses were performed in R (v3.6.2),³² with the significance threshold set at
226 false discovery rate (FDR) < 0.05. Variance components analysis was used to identify
227 influential variables for inclusion in DGE models (Fig. S2). Cell proportion estimates were
228 found to explain large fractions of expression variance, and adjusting for cell proportions
229 reduced the number of significant associations but improved consistency between drug
230 subgroups, with fewer highly significant modules showing significant drug-by-response
231 interaction effects compared to the unadjusted results (Fig. S3). As this study was not
232 designed to compare between drug subgroups, we focused on models adjusted for cell
233 composition, where the improved consistency allows us to pool expression data from both
234 subgroups for greater statistical power. For all DGE models, cell proportions, sequencing
235 batch, age of onset (the patient's age at disease diagnosis), disease duration, body mass

236 index (BMI), anti-TNF drug type, prior surgery, and smoking were included as fixed effects;
237 and patient was included as a random effect.

238

239 Per-gene linear mixed-effects models fit using DREAM (variancePartition v1.16.1)³³ were
240 used to detect pairwise DGE between study groups. Additionally, natural cubic splines
241 (splines::ns)³² were fit to explore non-linear expression trajectories over all four timepoints,
242 modelling expression as a function of study day in each group (first drug dose = study day 0).
243 Different expression trajectories were detected by testing for differences in spline
244 parameters between groups. Significant genes from the spline analysis were hierarchically
245 clustered by their mean expression in PR and PNR at each timepoint, and the gap statistic³⁴
246 was used to define clusters of genes with distinct trajectories. The spline analyses were only
247 performed with drug subgroups pooled, as relatively small sample sizes at week 30 and
248 week 54 precluded stratification by drug.

249

250 Rank-based gene set enrichment analyses (tmod::tmodCERNOtest, v0.46.2),³⁵ using blood
251 transcriptomic modules (BTMs) were performed to identify coordinately up or
252 downregulated gene sets. These modules are sets of genes that are coexpressed in whole
253 blood, derived by Li et al.³⁶ (module names prefixed with “LI”) and Chaussabel et al.³⁷
254 (prefixed “DC”) from publicly available expression datasets. Gene set overrepresentation
255 analyses were run for BTMs (tmod::tmodHGtest) and other publicly-available gene sets
256 (gprofiler2::gost, v0.2.0).³⁸

257

258 Single-sample gene set enrichment scores (ssGSEA,
259 <https://github.com/broadinstitute/ssGSEA2.0/>) were computed as a summary measure of

260 module expression in a sample, both at baseline and at week 14. Predictive models using
261 clinical variables, cell proportions, and module expression scores (baseline or week 14) to
262 predict response status were constructed using caret (v6.0-86).³⁹ Multiple predictive
263 algorithms were evaluated: penalised and regularised logistic regression methods, parallel
264 random forest, eXtreme Gradient Boosting, support vector machines with a radial basis, *k*-
265 nearest neighbours, naive Bayes, and Gaussian process models. Bootstrapping (50
266 replicates) with the AUC metric was used to tune models, evaluate internal performance,
267 and perform model selection. Pairwise tests for the difference in AUCs were performed with
268 pROC.⁴⁰

269 **Ethical statement**

270 The South West Research Ethics committee approved the study (Research Ethics Committee
271 reference: 12/SW/0323) in January 2013. Patients were included after providing informed,
272 written consent. The study is registered with ClinicalTrials.gov identifier NCT03088449, and
273 the protocol is available at <https://www.ibdresearch.co.uk/pants/>.

274

275 Results

276 Baseline expression modules are associated with post- 277 induction primary non-response

278 After RNA-seq quantification and quality control, expression data was available for 15 511
279 genes and 814 samples. These samples were from 324 patients, whose characteristics are
280 shown in Table 1. No single gene had a significant association between week 0 expression
281 and response in the infliximab subgroup (86 PR, 59 PNR). Expression of NK cell (LI.M7.2) and
282 T cell (LI.M7.1, LI.M7.0) gene modules (BTMs)³⁶ was significantly lower in responders.
283 Exclusively in the adalimumab subgroup (66 PR, 57 PNR), *PDIA5* (log2 fold change (FC) = -
284 0.3512, FDR = 0.006777), *KCNN3* (log2 FC = -0.8798, FDR = 0.006777), and *IGKV1-9* (log2 FC
285 = -1.223, FDR = 0.04518) had significantly lower expression in responders (Fig. 1a). This was
286 accompanied by lower expression of plasma cell/immunoglobulin (LI.M156.0, LI.M156.1)
287 and cell cycle (LI.M4.0, LI.M4.1) modules (Fig. 1b). This heterogeneity between drug
288 subgroups was robust to model form (Fig. S4) and differences in sample size between
289 subgroups (Fig. S5). A pooled analysis was performed to identify modules consistently
290 differentially expressed in both drug subgroups (152 PR, 116 PNR). MHC-TLR7-TLR8 cluster
291 (LI.M146), antigen presentation (LI.M71, LI.M95.0), and myeloid cell enriched receptor and
292 transporter (LI.M4.3) modules were found to be upregulated at baseline in responders. We
293 also examined previously reported baseline markers in gut mucosal biopsies and blood from
294 the literature,^{14,15,18,41} and did not find them to be significant in this study (Fig. 1a).

295 **Expression changes from baseline to post-induction are**
296 **largely amplified in primary responders**

297 To characterise the changes in gene expression induced by anti-TNF therapy, we compared
298 expression at baseline to expression post-induction, and also estimated the difference
299 between expression changes in responders and non-responders (the timepoint-by-response
300 interaction). As expression changes from week 0 to week 14 were relatively consistent
301 between patients on infliximab and adalimumab after adjusting for cell composition (Fig.
302 S6), we pooled drug subgroups for these models. We found that 5 572 and 626 genes were
303 differentially expressed in responders and non-responders respectively, with 179 genes
304 having a significant timepoint-by-response interaction. Of the genes differentially expressed
305 between week 14 and week 0 in both responders and non-responders, and with a significant
306 timepoint-by-response interaction, nearly all (31/32 genes) had an expression change that
307 was amplified in responders (Fig. 2a). For example, *CD177*, a neutrophil marker upregulated
308 during inflammation, was downregulated at week 14 in responders to a much greater
309 extent ($\log_2 \text{FC} = -2.225$, $\text{FDR} = 4.104 \times 10^{-17}$) than in non-responders ($\log_2 \text{FC} = -0.8981$, FDR
310 $= 0.004598$; interaction $\text{FDR} = 0.008247$). Modules differentially expressed between week 0
311 to week 14 included upregulation of B cell (LI.M47.0), plasma cell (LI.M156.0), and T cell
312 activation (LI.M7.1) modules; and downregulation of immune activation (LI.M37.0),
313 monocyte (LI.M11.0), neutrophil (LI.M37.1), and TLR and inflammatory signalling (LI.M16)
314 modules (Fig. 2b). Amplification of expression changes in responders was also observed at
315 the module level, with nearly all module expression changes aligned in the same direction in
316 responders and non-responders, but with larger effects in responders.
317

318 In contrast, *GBP2* (a member of a family of guanylate-binding proteins induced by
319 interferons⁴²) was downregulated from week 0 to week 14 in responders (\log_2 FC = -0.1783,
320 FDR = 0.004878), but upregulated in non-responders (\log_2 FC = 0.1849, FDR = 0.04502;
321 interaction FDR = 0.005977). At the module level, upregulation of type I interferon response
322 (LI.M127), activated dendritic cell (LI.M165), and antiviral IFN signature (LI.M75) modules
323 was observed in non-responders but not in responders (Fig. 2b). Further gene set
324 enrichment analyses using modules from Chaussabel et al.³⁷ also revealed annotated
325 interferon gene modules to be significantly upregulated at week 14 in non-responders:
326 DC.M3.4, containing *STAT2*, *GBP5*, and *PARP14* (FDR = 3.447×10^{-21}); as well two modules
327 containing *IFIT3* and *GBP2*, DC.M1.2 (FDR = 9.492×10^{-16}) and DC.M5.12 (FDR = 1.355×10^{-13}).
328 None of these modules were differentially expressed from week 0 to week 14 in
329 responders (Fig. 2c), suggesting upregulation of interferon pathways at week 14 occurs
330 uniquely in primary non-responders to anti-TNF.

331 **Sustained expression differences between responders and** 332 **non-responders during maintenance**

333 As PANTS was an observational study, it was possible to retain patients who continued with
334 anti-TNF therapy even after meeting the study definition of PNR at week 14, enabling us to
335 sample the blood transcriptome at week 30 and week 54 during the maintenance period.
336 Leveraging all 814 samples over the four study timepoints, we tested for general differences
337 in expression trajectory over time, detecting 210 differentially expressed genes between
338 responders and non-responders after adjustment for cell composition. To visualise the
339 expression of these genes and identify common patterns of expression change during anti-

340 TNF therapy, significant genes were hierarchically clustered by their expression. Six clusters
341 were identified (Fig. 3a), each with distinct expression trajectories for responders and non-
342 responders (Fig. 3b). Cluster 1 largely comprised genes previously found to have a significant
343 difference in expression change from week 0 to week 14 between responders and non-
344 responders (97/132 genes). The most significant gene was *KREMEN1* (FDR = 4.287×10^{-4}),
345 part of an inflammatory apoptotic pathway in gut epithelium.⁴³ Cluster 1 genes were
346 enriched in modules associated with myeloid cells and monocytes (LI.M81, hypergeometric
347 test, FDR = 2.115×10^{-6}), platelet activation (LI.M196, 1.348×10^{-5}), immune activation
348 (LI.M37.0, 1.436×10^{-4}), and TLR and inflammatory signalling (LI.M16, FDR = 2.365×10^{-3})
349 (Fig. 3c). Expression trajectories showed cluster 1 genes were more downregulated from
350 baseline in responders than non-responders, likely representing a lower inflammatory state
351 in responders by week 14 that is also maintained at week 30 and week 54. An opposing
352 trend was observed in cluster 5, which contained genes enriched for B cell
353 development/activation (LI.M58, FDR = 0.01653) that were more upregulated from baseline
354 in responders than non-responders.

355

356 Cluster 3 was uniquely enriched for type I interferon response genes (LI.M127, FDR =
357 0.005681, Fig. 3c). Subsequent enrichment analyses using publicly-available gene sets³⁸
358 revealed enrichments for type II interferon signalling (WP:WP619, adj. $p = 0.0002826$), and
359 for genes containing putative transcription factor (TF) binding motifs for interferon
360 regulatory factors IRF7 (TF:M00453_1, adj. $p = 0.004768$) and IRF8 (TF:M11684_1, adj. $p =$
361 0.006853; TF:M11685_1, adj. $p = 0.01136$) (Fig. 3d). The genes in cluster 3 showed opposing
362 directions of expression change from week 0 to week 14 in responders versus non-
363 responders, generating expression differences at week 14 that were sustained at week 30

364 and week 54. Of the nine genes in the cluster, eight genes (*STAT1*, *BATF2*, *GBP1*, *GBP5*, *IRF1*,
365 *TAP1*, *APOL1*, *APOL2*) had a significant interaction between week 0 to week 14 expression
366 change and response status in the per-gene differential expression analyses. *IRF1* and *STAT1*
367 are key transcription factors for interferon signalling.^{44,45} Unlike the majority of genes that
368 followed trajectories of greater expression change in responders, genes in interferon
369 response pathways were uniquely upregulated in non-responders.

370 **Prediction of primary non-response from gene expression** 371 **and clinical variables**

372 Given there were multiple significant module-level associations with primary response
373 status, we evaluated the ability to predict response from module expression, as well as cell
374 proportions and clinical variables. Single-sample gene set enrichment scores (ssGSEA) were
375 computed to summarise module expression per sample. Using baseline module scores, the
376 median resampling AUCs over all combinations of algorithms and predictor sets ranged
377 between 0.5541 and 0.6686 (Fig. S7). The best-performing model was regularised logistic
378 regression (regLogistic; hyperparameters: cost = 0.25, loss = L1, epsilon = 0.01) using clinical
379 variables, cell proportions, and module scores as predictors, giving a median resampling
380 AUC of 0.6686, a median sensitivity of 0.5392, and a median specificity of 0.6852, where
381 non-response was the positive class, with a prevalence of 43% (116/268). Including cell
382 proportions and module scores improved predictive performance compared to the model
383 using only clinical variables (bootstrap $p = 0.02629$), but the increase in AUC was only 2.5%
384 (Fig. 4a). This suggests that clinical variables provided the greatest contributions to baseline

385 prediction, especially those variables with high importance in all three models (high
386 absolute t -statistics): smoking history, BMI, and baseline steroid usage (Fig. 4b).
387
388 Performance for predicting primary non-response could be improved by using week 14 cell
389 proportions and module scores (median resampling AUC range 0.6216-0.7957, Fig. S8). The
390 prevalence of non-response at week 14 was 42% (104/246). The best performing predictor
391 was again regularised logistic regression incorporating clinical variables, cell proportions,
392 and module scores as predictors (median resampling AUC = 0.7957, sensitivity = 0.6248,
393 specificity = 0.7961). Here the addition of module scores to the predictor dataset had a
394 larger benefit on predictive performance, with a 6.5% increase in AUC comparing the full
395 model to the model including only clinical variables and cell proportions (bootstrap $p =$
396 2.306×10^{-10}), with a (Fig. 4c). The week 14 modules with the highest variable importance
397 included TLR and inflammatory signalling (LI.M16), chaperonin mediated protein folding
398 (LI.M204.0, LI.M204.1) and translation initiation factor 3 complex (LI.M245) modules (Fig.
399 4d). Greater predictive performance at week 14 likely reflects the larger clinical and
400 expression differences observed between responders and non-responders after the
401 induction period.

402

403 Discussion

404 We found substantial differences in whole blood gene expression between anti-TNF primary
405 responders and non-responders in the PANTS cohort. At baseline, three single-gene
406 associations detected in the adalimumab subgroup implicated similar cell types; *IGKV1-9*
407 encodes the immunoglobulin light chain variable region that forms part of antibodies
408 produced by plasma cells, *KCNN3* is annotated to a plasma cell surface signature module
409 from Li et al.³⁶ (LI.S3), and the expression of both *KCNN3* and *PDIA5* is high in plasma cells
410 (www.proteinatlas.org/humanproteome/immune+cell, v21.1) and positively correlated with
411 plasmablast frequencies in blood.⁴⁶ These genes were downregulated in responders, as was
412 the expression of plasma cell and immunoglobulin modules. In keeping with our
413 observations, Martin et al.¹⁷ identified plasma cells as part of a correlated module of cell
414 populations, where lower module expression in gut biopsies was associated with better
415 response to anti-TNFs. Baseline plasma cell abundances in gut biopsies have also been
416 reported to be lower in responders, albeit in relatively small cohorts of infliximab patients.¹⁶
417 Our findings lend credence that associations driven by immune cells observed in gut
418 biopsies may also be observable in blood, a more accessible tissue.

419

420 Previously reported single-gene baseline markers in gut biopsies and blood were non-
421 significant in this study. For example, *TREM1*^{16,18} was not significantly differentially
422 expressed between responders and non-responders in blood samples from PANTS patients.
423 Our observation is consistent with two recent trials of comparable sample size, SERENE-CD
424 and SERENE-UC, where baseline blood *TREM1* expression was not predictive of response in
425 either CD or UC patients.⁴⁷ A variety of factors could explain failures to replicate reported

426 markers from study to study. Many existing studies pool heterogeneous cohorts of patients
427 taking different anti-TNF drugs due to the scarcity of large datasets; even between arms of
428 the PANTS study, we observe within-study differences in expression. Additional between-
429 study variation arises from differences in clinical setting, tissues sampled (e.g. blood versus
430 gut biopsies), and definition of primary response (e.g. endoscopic versus clinical
431 parameters). Any two studies are unlikely to have adjusted for the same combinations of
432 covariates in modelling, including covariates such as cell composition that are influential for
433 bulk expression data. Finally, small sample sizes have considerable sampling error. Despite
434 only finding a small number of single-gene associations in PANTS, we detected multiple
435 novel module-level associations, such as baseline upregulation of MHC-mediated antigen
436 presentation modules in responders. We highly recommend the use of set-based methods
437 that draw on changes in multiple genes, as they are not only better powered, but may be
438 more reproducible compared to single-gene association tests.

439

440 Leveraging the longitudinal design of PANTS, we characterised the changes in blood gene
441 expression post-induction. Reduced expression of immune activation modules in responders
442 at week 14 is consistent with successful drug inhibition of TNF-mediated inflammation.
443 Decreased inflammation correlates with reduced neutrophil activation and reduced
444 monocyte recruitment,⁴⁸ and apoptosis of monocytes induced by anti-TNF in CD patients
445 has been previously described.⁴³ Certain B cell subsets are reduced in the blood of IBD
446 patients compared to controls,⁴⁹ so upregulation of B cell modules at week 14 may also
447 represent a shift towards health. Similar expression changes were observed in responders
448 and non-responders, but with greater magnitude in responders, potentially supporting the
449 hypothesis that response is a continuum. Gaujoux et al.¹⁶ found changes in cell proportions

450 in response to anti-TNF treatment were amplified in responders; here we demonstrate a
451 similar trend at the transcriptional level in PANTS. Post-induction expression differences
452 between responders and non-responders were sustained at week 30 and week 54 during
453 the anti-TNF maintenance period. Kennedy et al.⁸ found that “continuing standard dosing
454 regimens after primary non-response was rarely helpful” for inducing remission by week 54.
455 This phenomenon may also be reflected in the blood transcriptome, although non-
456 responders in the PANTS RNA-seq data were selected to exclude patients in remission by
457 week 54, so trajectories for non-responders at week 14 that eventually achieved remission
458 could not be observed. Thus, the observed difference in trajectories between responders
459 and non-responders is likely exaggerated.

460

461 Unlike the majority of baseline versus post-induction associations, expression changes in
462 genes and modules in the interferon pathway were uniquely upregulated in PANTS non-
463 responders. Previous studies in IBD are conflicting, with Samie et al.⁵⁰ reporting elevated
464 expression of interferon pathway genes in colonic biopsies from non-responders compared
465 to responders, with no significant change pre- versus post-treatment ($n \approx 40$); and
466 Mavragani et al.²⁰ reporting a post-treatment reduction in blood interferon expression only
467 in non-responders ($n = 30$). In studies of rheumatoid arthritis (RA), another IMID with
468 licensed anti-TNF therapies, increases in type I interferon-regulated gene expression in
469 blood after infliximab treatment were associated with poor clinical response (discovery $n =$
470 15, validation $n = 18$).⁵¹ A systematic review of our study with other studies reporting similar
471 associations between interferon pathway genes and anti-TNF response would not only help
472 resolve the direction of effect, if any, but provide an opportunity to consider the shared
473 biology of anti-TNF response in IBD, RA, and other immune-mediated diseases.

474

475 An important limitation of our analyses is that PANTS was not designed to compare
476 between drug subgroups. Differences between patients on different anti-TNF drugs can
477 arise from patient and physician preference, influenced by cost, disease severity, location,
478 and comorbidities. Unsurprisingly, many associations with response had significantly
479 different effect sizes in infliximab and adalimumab patient subgroups. We found that
480 adjusting DGE models for estimated proportions of major cell types as a proxy for these
481 uncontrolled factors alleviated heterogeneity between subgroups. However, the adjustment
482 is unlikely to work well for rare cell types, thus the associations we report may reflect
483 differences in cell proportions rather than per-cell expression. Given the myriad of other
484 factors that could drive the remaining heterogeneity, we strongly caution against
485 interpreting associations with different effects in the PANTS infliximab and adalimumab
486 subgroups as drug-driven differences with biological significance, and recommend that
487 future transcriptomic studies consider influential factors such as cell composition.

488

489 We were unable to build clinically useful predictive models of response incorporating
490 expression data. Using only baseline clinical variables, Kennedy et al.⁸ used logistic
491 regression with stepwise variable selection based on Akaike information criterion (AIC) to
492 predict response in the full PANTS cohort, achieving AUCs of 0.53 (95% CI 0.46–0.59) for
493 infliximab patients and 0.54 (0.46–0.62) for adalimumab patients. Whilst our best-
494 performing baseline model represents an improvement, expression data only contributed a
495 small amount of predictive power on top of clinical variables and cell composition.
496 Unsurprisingly, models had greater predictive power when provided with week 14
497 expression and cell composition data, and adding expression data also provided a

498 comparatively large increase in AUC. This suggests that when expression differences
499 between responders and non-responders are sufficiently large, transcriptomic markers do
500 provide unique information, and are not simply proxies for clinical variables or coarse
501 estimates of blood cell composition. A potential route to more effective prediction is to
502 consider whether expression differences arising early in the induction period can
503 discriminate between responders and non-responders. For example, Mesko et al.⁵² found
504 that week 2 blood gene expression was predictive of infliximab response in CD (discovery n
505 = 20, validation $n = 20$) and RA (discovery $n = 19$, validation $n = 15$) patients. More recently,
506 Mishra et al.⁵³ trained random forest models using blood DNA methylation and gene
507 expression measured in IBD patients receiving infliximab ($n = 37$). They did not find
508 consistent baseline-only predictive signatures, but a model combining baseline with week 2
509 measurements predicted response in the Mesko et al. cohort with 85% accuracy (95% CI:
510 62–97%). As we observed in PANTS, expression differences between responders and non-
511 responders were far greater by week 14 than at baseline. Post-induction associations were
512 also more consistent between drug subgroups, as baseline differences are diluted by the
513 large transcriptomic perturbation from taking an anti-TNF. Expression changes in the innate
514 immune system are observable within hours of treatment initiation,⁵³ and robust prediction
515 of non-response within that timeframe may be more valuable than a less reliable prediction
516 at baseline.

517

518 In conclusion, we observed significant differences in gene module expression between
519 responders and non-responders to anti-TNF therapy in the whole blood of PANTS CD
520 patients at baseline and post-treatment timepoints. Interferon-induced genes were
521 uniquely upregulated post-induction in non-responders, going against the general trend of

522 amplified transcriptomic change in responders versus non-responders. However, there is
523 not yet a consensus on the direction of these effects, and we were unable to robustly
524 predict response with our current sample size. To obtain accurate prediction, especially
525 from baseline data, leveraging large upcoming datasets with paired drug response
526 phenotypes and transcriptomic data such as the 1000IBD project will be essential.⁵⁴

527

528 **Funding**

529 PANTS is an investigator-led study funded by CORE (renamed Guts UK in 2018), the research
530 charity of the British Society of Gastroenterology, and by unrestricted educational grants
531 from AbbVie Inc, USA, Merck Sharp & Dohme Ltd, UK, NAPP Pharmaceuticals Ltd, UK, Pfizer
532 Ltd, USA, Celltrion Healthcare, South Korea, and Cure Crohn's Colitis (Scottish IBD Charity).
533 The sponsor of the PANTS study is the Royal Devon and Exeter National Health Service
534 Foundation Trust. Data generation, sequencing, and sample analysis were funded by
535 AbbVie. This research was funded in whole, or in part, by the Wellcome Trust [grant
536 numbers 206194, 108413/A/15/D, and 222850/Z/21/Z]. For the purpose of Open Access,
537 the author has applied a CC BY public copyright licence to any Author Accepted Manuscript
538 version arising from this submission.

539 **Conflicts of Interest**

540 Mark Reppell, Nizar Smaoui, Jeffery F. Waring, Valerie Pivorunas, and Heath Guay are
541 employees of AbbVie and may own stock and/or options. Simeng Lin reports non-financial
542 support from Pfizer outside the submitted work. James R. Goodhand reports grants from F.
543 Hoffmann-La Roche AG, grants from Biogen Inc, grants from Celltrion Healthcare, grants
544 from Galapagos NV and non-financial support from Immundiagnostik outside the conduct of
545 the study. Nicholas A. Kennedy reports grants from F. Hoffmann-La Roche AG, grants from
546 Biogen Inc, grants from Celltrion Healthcare, grants from Galapagos NV and non-financial
547 support from Immundiagnostik; grants and non-financial support from AbbVie, grants and

548 personal fees from Celltrion, personal fees and non-financial support from Janssen, personal
549 fees from Takeda, and personal fees and non-financial support from Dr Falk, outside the
550 submitted work. Tariq Ahmad reports grants and non-financial support from F. Hoffmann-La
551 Roche AG, grants from Biogen Inc, grants from Celltrion Healthcare, grants from Galapagos
552 NV and non-financial support from Immundiagnostik; personal fees from Biogen inc, grants
553 and personal fees from Celltrion Healthcare, personal fees and non-financial support from
554 Immundiagnostik, personal fees from Takeda, personal fees from ARENA, personal fees
555 from Gilead, personal fees from Adcock Ingram Healthcare, personal fees from Pfizer,
556 personal fees from Genentech and non-financial support from Tillotts, outside the
557 submitted work. Carl A. Anderson has received consultancy or lectureship fees from
558 Genomics plc, BridgeBio, and GSK. The remaining authors have no conflicts of interest to
559 report.

560 **Author Contributions**

561 SL, NC, JRG, NAK, TA, and CAA participated in the conception and design of the study. CB
562 was the project manager and coordinated recruitment. BYHB, MR, NS, JFW, VP, HG, SL, NC,
563 CB, JRG, NAK, TA, and CAA were involved in the acquisition, analysis, or interpretation of
564 data. The data analysis was performed by BYHB, MR, SL, and NAK. Drafting of the
565 manuscript was conducted by BYHB. All authors contributed to the revision of the
566 manuscript for critically important intellectual content. TA obtained the funding for the
567 study. CAA is the guarantor of the manuscript. All authors contributed to the final approval
568 of the manuscript.

569 **Acknowledgements**

570 We would like to acknowledge Jim Butler, Areej Ammar, Erin Murphy, Stephen Abel,
571 Elizabeth Asque, and Justin Wade Davis from AbbVie Inc., Chicago, IL, USA for supporting
572 this work. Laboratory tests were undertaken by the Exeter Blood Sciences Laboratory at the
573 Royal Devon & Exeter (RD&E) NHS Trust (<https://www.exeterlaboratory.com/>). The Exeter
574 NIHR Clinical Research Facility coordinated sample storage and management. The sponsor
575 of the study was the Royal Devon and Exeter NHS Foundation Trust. Simeng Lin is supported
576 by a Wellcome GW4-CAT fellowship (222850/Z/21/Z). Neil Chanchlani acknowledges
577 support from Crohn's & Colitis UK. We thank the patients who participated in the PANTS
578 study, the Inflammatory Bowel Disease Pharmacogenetic Study Group, and research nurses
579 who collected clinical data and biological samples at each study visit. We acknowledge the
580 study co-ordinators of the Exeter IBD Research Group, United Kingdom: Marian Parkinson
581 and Helen Gardner-Thorpe for their ongoing administrative support to the study.

582 **Data Availability**

583 Individual participant de-identified data that underlie the results reported in this article will
584 be available immediately after publication for a period of 5 years. The data will be made
585 available to investigators whose proposed use of the data has been approved by an
586 independent review committee. Analyses will be restricted to the aims in the approved
587 proposal. Proposals should be directed to Tariq Ahmad (tariq.ahmad1@nhs.net) and Carl A.

588 Anderson (ca3@sanger.ac.uk). To gain access data requestors will need to sign a data access

589 agreement.

590

591 **References**

- 592 1. Adegbola, S.O., Sahnan, K., Warusavitarne, J., Hart, A., and Tozer, P. (2018). Anti-TNF
593 Therapy in Crohn's Disease. *Int. J. Mol. Sci.* *19*, 2244. [10.3390/ijms19082244](https://doi.org/10.3390/ijms19082244).
- 594 2. Lichtenstein, G.R. (2013). Comprehensive review: antitumor necrosis factor agents in
595 inflammatory bowel disease and factors implicated in treatment response. *Ther. Adv.*
596 *Gastroenterol.* *6*, 269–293. [10.1177/1756283X13479826](https://doi.org/10.1177/1756283X13479826).
- 597 3. Neurath, M.F. (2014). New targets for mucosal healing and therapy in inflammatory
598 bowel diseases. *Mucosal Immunol.* *7*, 6–19. [10.1038/mi.2013.73](https://doi.org/10.1038/mi.2013.73).
- 599 4. Levin, A.D., Wildenberg, M.E., and van den Brink, G.R. (2016). Mechanism of Action of
600 Anti-TNF Therapy in Inflammatory Bowel Disease. *J. Crohns Colitis* *10*, 989–997.
601 [10.1093/ecco-jcc/jjw053](https://doi.org/10.1093/ecco-jcc/jjw053).
- 602 5. Roda, G., Jharap, B., Neeraj, N., and Colombel, J.-F. (2016). Loss of Response to Anti-TNFs:
603 Definition, Epidemiology, and Management: *Clin. Transl. Gastroenterol.* *7*, e135.
604 [10.1038/ctg.2015.63](https://doi.org/10.1038/ctg.2015.63).
- 605 6. Ben-Horin, S., Kopylov, U., and Chowers, Y. (2014). Optimizing anti-TNF treatments in
606 inflammatory bowel disease. *Autoimmun. Rev.* *13*, 24–30. [10.1016/j.autrev.2013.06.002](https://doi.org/10.1016/j.autrev.2013.06.002).
- 607 7. Flamant, M., and Roblin, X. (2018). Inflammatory bowel disease: towards a personalized
608 medicine. *Ther. Adv. Gastroenterol.* *11*, 1756283X1774502. [10.1177/1756283X17745029](https://doi.org/10.1177/1756283X17745029).
- 609 8. Kennedy, N.A., Heap, G.A., Green, H.D., Hamilton, B., Bewshea, C., Walker, G.J., Thomas,
610 A., Nice, R., Perry, M.H., Bouri, S., et al. (2019). Predictors of anti-TNF treatment failure in

- 611 anti-TNF-naive patients with active luminal Crohn's disease: a prospective, multicentre,
612 cohort study. *Lancet Gastroenterol. Hepatol.* 4, 341–353. 10.1016/S2468-1253(19)30012-
613 3.
- 614 9. D'Haens, G.R., Panaccione, R., Higgins, P.D.R., Vermeire, S., Gassull, M., Chowers, Y.,
615 Hanauer, S.B., Herfarth, H., Hommes, D.W., Kamm, M., et al. (2011). The London Position
616 Statement of the World Congress of Gastroenterology on Biological Therapy for IBD With
617 the European Crohn's and Colitis Organization: When to Start, When to Stop, Which Drug
618 to Choose, and How to Predict Response?. *Am. J. Gastroenterol.* 106, 199–212.
619 10.1038/ajg.2010.392.
- 620 10. Ding, N.S., Hart, A., and De Cruz, P. (2016). Systematic review: predicting and optimising
621 response to anti-TNF therapy in Crohn's disease - algorithm for practical management.
622 *Aliment. Pharmacol. Ther.* 43, 30–51. 10.1111/apt.13445.
- 623 11. Kopylov, U., and Seidman, E. (2016). Predicting durable response or resistance to
624 antitumor necrosis factor therapy in inflammatory bowel disease. *Ther. Adv.*
625 *Gastroenterol.* 9, 513–526. 10.1177/1756283X16638833.
- 626 12. Digby-Bell, J.L., Atreya, R., Monteleone, G., and Powell, N. (2019). Interrogating host
627 immunity to predict treatment response in inflammatory bowel disease. *Nat. Rev.*
628 *Gastroenterol. Hepatol.* 10.1038/s41575-019-0228-5.
- 629 13. Noor, N.M., Verstockt, B., Parkes, M., and Lee, J.C. (2020). Personalised medicine in
630 Crohn's disease. *Lancet Gastroenterol. Hepatol.* 5, 80–92. 10.1016/S2468-
631 1253(19)30340-1.

- 632 14. Arijs, I., Li, K., Toedter, G., Quintens, R., Van Lommel, L., Van Steen, K., Leemans, P., De
633 Hertogh, G., Lemaire, K., Ferrante, M., et al. (2009). Mucosal gene signatures to predict
634 response to infliximab in patients with ulcerative colitis. *Gut* 58, 1612–1619.
635 10.1136/gut.2009.178665.
- 636 15. Arijs, I., Quintens, R., Van Lommel, L., Van Steen, K., De Hertogh, G., Lemaire, K.,
637 Schraenen, A., Perrier, C., Van Assche, G., Vermeire, S., et al. (2010). Predictive value of
638 epithelial gene expression profiles for response to infliximab in Crohn’s disease. *Inflamm.*
639 *Bowel Dis.* 16, 2090–2098. 10.1002/ibd.21301.
- 640 16. Gaujoux, R., Starosvetsky, E., Maimon, N., Vallania, F., Bar-Yoseph, H., Pressman, S.,
641 Weisshof, R., Goren, I., Rabinowitz, K., Waterman, M., et al. (2019). Cell-centred meta-
642 analysis reveals baseline predictors of anti-TNF α non-response in biopsy and blood of
643 patients with IBD. *Gut* 68, 604–614. 10.1136/gutjnl-2017-315494.
- 644 17. Martin, J.C., Chang, C., Boschetti, G., Ungaro, R., Giri, M., Grout, J.A., Gettler, K., Chuang,
645 L., Nayar, S., Greenstein, A.J., et al. (2019). Single-Cell Analysis of Crohn’s Disease Lesions
646 Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy.
647 *Cell* 178, 1493-1508.e20. 10.1016/j.cell.2019.08.008.
- 648 18. Verstockt, B., Verstockt, S., Dehairs, J., Ballet, V., Blevi, H., Wollants, W.-J., Breynaert, C.,
649 Van Assche, G., Vermeire, S., and Ferrante, M. (2019). Low TREM1 expression in whole
650 blood predicts anti-TNF response in inflammatory bowel disease. *EBioMedicine* 40, 733–
651 742. 10.1016/j.ebiom.2019.01.027.
- 652 19. West, N.R., Hegazy, A.N., Owens, B.M.J., Bullers, S.J., Linggi, B., Buonocore, S., Coccia,
653 M., Görtz, D., This, S., Stockenhuber, K., et al. (2017). Oncostatin M drives intestinal

- 654 inflammation and predicts response to tumor necrosis factor–neutralizing therapy in
655 patients with inflammatory bowel disease. *Nat. Med.* 23, 579–589. 10.1038/nm.4307.
- 656 20. Mavragani, C.P., Nezos, A., Dovrolis, N., Andreou, N.-P., Legaki, E., Sechi, L.A., Bamias,
657 G., and Gazouli, M. (2020). Type I and II Interferon Signatures Can Predict the Response
658 to Anti-TNF Agents in Inflammatory Bowel Disease Patients: Involvement of the
659 Microbiota. *Inflamm. Bowel Dis.* 26, 1543–1553. 10.1093/ibd/izaa216.
- 660 21. Verstockt, B., Verstockt, S., Blevi, H., Cleynen, I., de Bruyn, M., Van Assche, G., Vermeire,
661 S., and Ferrante, M. (2019). TREM-1, the ideal predictive biomarker for endoscopic
662 healing in anti-TNF-treated Crohn’s disease patients? *Gut* 68, 1531–1533. 10.1136/gutjnl-
663 2018-316845.
- 664 22. Broad Institute (2019). Picard toolkit. <https://broadinstitute.github.io/picard/>.
- 665 23. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,
666 M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*
667 29, 15–21. 10.1093/bioinformatics/bts635.
- 668 24. Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in
669 Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* 27, 491–
670 499. 10.1101/gr.209601.116.
- 671 25. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose
672 program for assigning sequence reads to genomic features. *Bioinforma. Oxf. Engl.* 30,
673 923–930. 10.1093/bioinformatics/btt656.

- 674 26. Piasecka, B., Duffy, D., Urrutia, A., Quach, H., Patin, E., Posseme, C., Bergstedt, J.,
675 Charbit, B., Rouilly, V., MacPherson, C.R., et al. (2018). Distinctive roles of age, sex, and
676 genetics in shaping transcriptional variation of human immune responses to microbial
677 challenges. *Proc. Natl. Acad. Sci.* *115*, E488–E497. [10.1073/pnas.1714765115](https://doi.org/10.1073/pnas.1714765115).
- 678 27. Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson,
679 H.H., Wiencke, J.K., and Kelsey, K.T. (2012). DNA methylation arrays as surrogate
680 measures of cell mixture distribution. *BMC Bioinformatics* *13*, 86. [10.1186/1471-2105-13-](https://doi.org/10.1186/1471-2105-13-86)
681 [86](https://doi.org/10.1186/1471-2105-13-86).
- 682 28. Lin, S., Hannon, E., Reppell, M., Waring, J. F., Smaoui, N., Pivorunas, V., Guay, H.,
683 Chanchlani, N., Bewshea, C., Bai, B. Y. H., Kennedy, N. A., Goodhand, J. R., Mill, J., Ahmad,
684 T., & PANTS Consortium. (2023). Whole blood DNA methylation changes are associated
685 with anti-TNF drug concentration in patients with Crohn’s disease. *MedRxiv*.
686 <https://doi.org/10.1101/2023.03.22.23287574>
- 687 29. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package
688 for differential expression analysis of digital gene expression data. *Bioinformatics* *26*,
689 139–140. [10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616).
- 690 30. Hoffman, G.E., and Schadt, E.E. (2016). variancePartition: interpreting drivers of
691 variation in complex gene expression studies. *BMC Bioinformatics* *17*. [10.1186/s12859-](https://doi.org/10.1186/s12859-016-1323-z)
692 [016-1323-z](https://doi.org/10.1186/s12859-016-1323-z).
- 693 31. Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: precision weights unlock
694 linear model analysis tools for RNA-seq read counts. *Genome Biol.* *15*, 1–17.

- 695 32. R Core Team (2019). R: A language and environment for statistical computing.
- 696 33. Hoffman, G.E., and Roussos, P. (2020). dream: Powerful differential expression analysis
697 for repeated measures designs. *Bioinformatics*. 10.1093/bioinformatics/btaa687.
- 698 34. Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a
699 data set via the gap statistic. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 63, 411–423.
700 10.1111/1467-9868.00293.
- 701 35. Weiner 3rd, J., and Domaszewska, T. (2016). tmod: an R package for general and
702 multivariate enrichment analysis. 10.7287/peerj.preprints.2420v1.
- 703 36. Li, S., Roupahel, N., Duraisingham, S., Romero-Steiner, S., Presnell, S., Davis, C., Schmidt,
704 D.S., Johnson, S.E., Milton, A., Rajam, G., et al. (2013). Molecular signatures of antibody
705 responses derived from a systems biology study of five human vaccines. *Nat. Immunol.*
706 15, 195–204. 10.1038/ni.2789.
- 707 37. Chaussabel, D., Quinn, C., Shen, J., Patel, P., Glaser, C., Baldwin, N., Stichweh, D.,
708 Blankenship, D., Li, L., Munagala, I., et al. (2008). A Modular Analysis Framework for
709 Blood Genomics Studies: Application to Systemic Lupus Erythematosus. *Immunity* 29,
710 150–164. 10.1016/j.immuni.2008.05.012.
- 711 38. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019).
712 g:Profiler: a web server for functional enrichment analysis and conversions of gene lists
713 (2019 update). *Nucleic Acids Res.* 47, W191–W198. 10.1093/nar/gkz369.
- 714 39. Kuhn, M. (2020). caret: Classification and regression training.

- 715 40. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., and Müller, M.
716 (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves.
717 *BMC Bioinformatics* 12, 77. 10.1186/1471-2105-12-77.
- 718 41. Salvador-Martín, S., Raposo-Gutiérrez, I., Navas-López, V.M., Gallego-Fernández, C.,
719 Moreno-Álvarez, A., Solar-Boga, A., Muñoz-Codoceo, R., Magallares, L., Martínez-Ojinaga,
720 E., Fobelo, M.J., et al. (2020). Gene Signatures of Early Response to Anti-TNF Drugs in
721 Pediatric Inflammatory Bowel Disease. *Int. J. Mol. Sci.* 21, 3364. 10.3390/ijms21093364.
- 722 42. Tretina, K., Park, E.-S., Maminska, A., and MacMicking, J.D. (2019). Interferon-induced
723 guanylate-binding proteins: Guardians of host defense in health and disease. *J. Exp. Med.*
724 216, 482–500. 10.1084/jem.20182031.
- 725 43. Lügering, A., Schmidt, M., Lügering, N., Pauels, H.-G., Domschke, W., and Kucharzik, T.
726 (2001). Infliximab induces apoptosis in monocytes from patients with chronic active
727 Crohn's disease by using a caspase-dependent pathway. *Gastroenterology* 121, 1145–
728 1157. 10.1053/gast.2001.28702.
- 729 44. Liu, S.-Y., Sanchez, D.J., Aliyari, R., Lu, S., and Cheng, G. (2012). Systematic identification
730 of type I and type II interferon-induced antiviral factors. *Proc. Natl. Acad. Sci.* 109, 4239–
731 4244. 10.1073/pnas.1114981109.
- 732 45. Schneider, W.M., Chevillotte, M.D., and Rice, C.M. (2014). Interferon-Stimulated Genes:
733 A Complex Web of Host Defenses. *Annu. Rev. Immunol.* 32, 513–545. 10.1146/annurev-
734 immunol-032713-120231.

- 735 46. Tsang, J.S., Schwartzberg, P.L., Kotliarov, Y., Biancotto, A., Xie, Z., Germain, R.N., Wang,
736 E., Olnes, M.J., Narayanan, M., Golding, H., et al. (2014). Global Analyses of Human
737 Immune Variation Reveal Baseline Predictors of Postvaccination Responses. *Cell* *157*,
738 499–513. [10.1016/j.cell.2014.03.031](https://doi.org/10.1016/j.cell.2014.03.031).
- 739 47. Verstockt, B., Pivorunas, V., Al Mahi, N., Smaoui, N., Guay, H., Kennedy, N.A., Ahmad, T.,
740 Goodhand, J.R., Lin, S., Bai, B.Y.H., et al. (2022). Baseline Whole Blood Gene Expression of
741 *TREM-1* Does Not Predict Response to Adalimumab Treatment in Patients with Ulcerative
742 Colitis or Crohn’s Disease in the SERENE Studies. *Manuscr. Prep.*
- 743 48. Prame Kumar, K., Nicholls, A.J., and Wong, C.H.Y. (2018). Partners in crime: neutrophils
744 and monocytes/macrophages in inflammation and disease. *Cell Tissue Res.* *371*, 551–565.
745 [10.1007/s00441-017-2753-2](https://doi.org/10.1007/s00441-017-2753-2).
- 746 49. Pararasa, C., Zhang, N., Tull, T.J., Chong, M.H.A., Siu, J.H.Y., Guesdon, W., Chavele, K.M.,
747 Sanderson, J.D., Langmead, L., Kok, K., et al. (2019). Reduced CD27-IgD- B Cells in Blood
748 and Raised CD27-IgD- B Cells in Gut-Associated Lymphoid Tissue in Inflammatory Bowel
749 Disease. *Front. Immunol.* *10*. [10.3389/fimmu.2019.00361](https://doi.org/10.3389/fimmu.2019.00361).
- 750 50. Samie, M., Lim, J., Verschueren, E., Baughman, J.M., Peng, I., Wong, A., Kwon, Y.,
751 Senbabaoglu, Y., Hackney, J.A., Keir, M., et al. (2018). Selective autophagy of the adaptor
752 TRIF regulates innate inflammatory signaling. *Nat. Immunol.* *19*, 246–254.
753 [10.1038/s41590-017-0042-6](https://doi.org/10.1038/s41590-017-0042-6).
- 754 51. van Baarsen, L.G., Wijbrandts, C.A., Rustenburg, F., Cantaert, T., van der Pouw Kraan,
755 T.C., Baeten, D.L., Dijkmans, B.A., Tak, P.P., and Verweij, C.L. (2010). Regulation of IFN

- 756 response gene activity during infliximab treatment in rheumatoid arthritis is associated
757 with clinical response to treatment. *Arthritis Res. Ther.* **12**, R11. 10.1186/ar2912.
- 758 52. Mesko, B., Poliska, S., Vánca, A., Szekanecz, Z., Palatka, K., Hollo, Z., Horvath, A.,
759 Steiner, L., Zahuczky, G., Podani, J., et al. (2013). Peripheral blood derived gene panels
760 predict response to infliximab in rheumatoid arthritis and Crohn's disease. *Genome Med.*
761 **5**, 59. 10.1186/gm463.
- 762 53. Mishra, N., Aden, K., Blase, J.I., Baran, N., Bordoni, D., Tran, F., Conrad, C., Avalos, D.,
763 Jaeckel, C., Scherer, M., et al. (2022). Longitudinal multi-omics analysis identifies early
764 blood-based predictors of anti-TNF therapy response in inflammatory bowel disease.
765 *Genome Med.* **14**, 110. 10.1186/s13073-022-01112-z.
- 766 54. Imhann, F., Van der Velde, K.J., Barbieri, R., Alberts, R., Voskuil, M.D., Vich Vila, A., Collij,
767 V., Spekhorst, L.M., Van der Sloot, K.W.J., Peters, V., et al. (2019). The 1000IBD project:
768 multi-omics data of 1000 inflammatory bowel disease patients; data release 1. *BMC*
769 *Gastroenterol.* **19**. 10.1186/s12876-018-0917-5.
- 770 55. Sazonovs, A., Kennedy, N.A., Moutsianas, L., Heap, G.A., Rice, D.L., Reppell, M.,
771 Bewshea, C.M., Chanchlani, N., Walker, G.J., Perry, M.H., et al. (2019). HLA-DQA1*05
772 Carriage Associated With Development of Anti-Drug Antibodies to Infliximab and
773 Adalimumab in Patients With Crohn's Disease. *Gastroenterology*.
774 10.1053/j.gastro.2019.09.041.
- 775

776 **Figure Captions**

777 **Figure 1. Baseline expression associated with primary response, adjusted for cell**
778 **composition.**

779 a. Volcano plots of DGE between responders (PR) and non-responders (PNR) at week 0;
780 for infliximab (IFX), adalimumab (ADA), or with drug subgroups pooled. Annotated
781 genes show significant associations from this study, and previously reported
782 associations from the literature in both blood and gut biopsies. Dashed line shows
783 significance threshold at FDR = 0.05.

784 b. Top gene modules differentially expressed between PR and PNR at week 0. Columns
785 correspond to results for infliximab (IFX), adalimumab (ADA), difference between IFX
786 and ADA (IFX – ADA i.e. drug-by-response interaction), and pooled drug analyses.
787 The top 30 modules ranked by minimum FDR in any column are shown. Dashed lines
788 show significance thresholds at FDR = 0.05.

789

790 **Figure 2. Expression changes from baseline to post-induction are amplified in responders.**

791 a. Expression log₂ fold changes from week 0 to week 14 in primary responders (PR) and
792 non-responders (PNR), for genes that differentially expressed from week 0 to week
793 14 in both responders and non-responders, with a significantly different effect size
794 between responders and non-responders (top 10 labelled). The identity line is shown
795 by the dashed line.

796 b. Top modules differentially expressed between week 14 and week 0, adjusted for cell
797 composition. Columns show effects in primary responders (PR), non-responders
798 (PNR), and the primary responder minus non-responder difference. The top 30

799 modules ranked by minimum FDR in any column are shown. Vertical dashed line
800 shows significance threshold at $FDR = 0.05$.

801 c. Barcode plots showing interferon modules (LI/DC) specifically upregulated from
802 week 0 to week 14 in primary non-responders (PNR), but not in primary responders
803 (PR). Genes were ranked in ascending order by week 14 versus week 0 DGE z-
804 statistic. Curves show the cumulative fraction of genes in each module. Effect sizes
805 are represented by the area under the curves (AUCs). Diagonal line shows the null of
806 randomly-distributed ranks.

807

808 **Figure 3. Sustained expression differences between responders and non-responders**

809 **during maintenance.**

- 810 a. Gap statistic versus cluster number k from hierarchical clustering of genes with
811 significant expression differences between PR and PNR over all timepoints. Error
812 bars derived from 500 bootstraps. The optimal cluster number was selected to be $k =$
813 6 by the `factoextra::fviz_nbclust` `firstSEmax` criteria
814 (<https://rpkgs.datanovia.com/factoextra/index.html>).
- 815 b. Normalised expression over study timepoints for genes in each cluster. 95%
816 confidence intervals for expression are shown for each group at each timepoint.
- 817 c. Gene modules enriched in each cluster from gene set overrepresentation analyses
818 using `tmod::tmodHGtest`. Modules significantly enriched in any cluster are shown.
819 Vertical dashed line shows significance threshold at $FDR = 0.05$.
- 820 d. Gene sets enriched in cluster 3 from gene set overrepresentation analyses using
821 `gprofiler2::gost`. Vertical dashed line shows significance threshold at adjusted p-
822 value = 0.05 (`gost` `g:SCS` multiple testing correction method).

823

824 **Figure 4. Prediction of response status from clinical variables, cell proportions, and**
825 **expression data.**

826 ROC curves for the regLogistic method trained on each predictor dataset are shown at

827 baseline (a) and week 14 (c). ROC curves were plotted after merging all 50 resamples.

828 Primary non-response was used as the positive class. DeLong 95% confidence intervals for

829 the AUC are shown.

830 The ten most important variables from models trained on each predictor dataset are shown

831 for baseline (b) and week 14 (d) models. The overall variable importance score is computed

832 from the absolute value of the t -statistic for each predictor from the final tuned models.

833 Missing bars denote variable that were not in the predictor dataset for that model.

834 **Tables**

835 **Table 1. Baseline patient characteristics for the PANTS RNA-seq subcohort.**

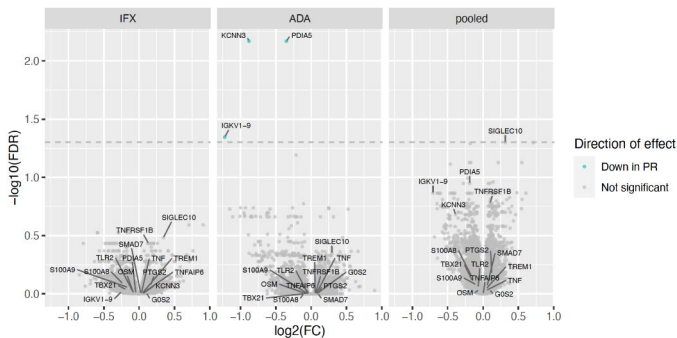
836 Characteristics are stratified by drug subgroup. Values shown are count and percentage for
 837 categorical variables, and mean and standard deviation for continuous variables. Nominal *p*-
 838 values are reported for the comparison between drug subgroups.

	adalimumab (ADA)	infliximab (IFX)	drugs pooled	<i>p</i> -value
Sex				0.317
(Col %)				Fisher exact
FEMALE	78 (48.4%)	89 (54.6%)	167 (51.5%)	
MALE	83 (51.6%)	74 (45.4%)	157 (48.5%)	
Age of onset (years)				0.774
Mean (SD)	33.3 (15.4)	32.8 (15.3)	33.1 (15.3)	Wilcoxon rank-sum
Missing	0	0	0	
Disease duration (years)				0.546
Mean (SD)	6.1 (8.1)	5.9 (7.7)	6.0 (7.9)	Wilcoxon rank-sum
Missing	0	0	0	
Smoking status				0.263
(Col %)				Fisher exact
Current	28 (17.4%)	36 (22.1%)	64 (19.8%)	
Ex	55 (34.2%)	43 (26.4%)	98 (30.2%)	
Never	78 (48.4%)	84 (51.5%)	162 (50.0%)	
Crohn's-related surgery				0.549
(Col %)				Fisher exact
NO	114 (70.8%)	110 (67.5%)	224 (69.1%)	
YES	47 (29.2%)	53 (32.5%)	100 (30.9%)	
On immunomodulator ever				0.543
(Col %)				Fisher exact
NO	23 (14.3%)	28 (17.2%)	51 (15.7%)	
YES	138 (85.7%)	135 (82.8%)	273 (84.3%)	
On immunomodulator at baseline				0.912
(Col %)				Fisher exact
NO	79 (49.1%)	81 (49.7%)	160 (49.4%)	
YES	82 (50.9%)	82 (50.3%)	164 (50.6%)	
On corticosteroids at baseline				0.011
(Col %)				Fisher exact
NO	113 (70.2%)	92 (56.4%)	205 (63.3%)	
YES	48 (29.8%)	71 (43.6%)	119 (36.7%)	
Baseline BMI				0.237
Mean (SD)	25.2 (6.2)	24.3 (5.5)	24.8 (5.9)	Wilcoxon rank-sum
Missing	0	0	0	
Primary response status				0.263
(Col %)				Fisher exact
Primary non-response	76 (47.2%)	66 (40.5%)	142 (43.8%)	
Primary response	85 (52.8%)	97 (59.5%)	182 (56.2%)	
CD8⁺ T cell (%)				0.380
Mean (SD)	2.8 (4.2)	2.8 (5.2)	2.8 (4.7)	Wilcoxon rank-sum
Missing	38	18	56	
CD4⁺ T cell (%s)				0.752
Mean (SD)	9.2 (6.3)	9.2 (6.8)	9.2 (6.5)	Wilcoxon rank-sum
Missing	38	18	56	
B cell (%s)				0.094
Mean (SD)	1.9 (2.0)	1.5 (1.9)	1.7 (1.9)	Wilcoxon rank-sum
Missing	38	18	56	
Monocyte (%s)				0.497
Mean (SD)	8.9 (3.5)	9.2 (3.7)	9.0 (3.6)	Wilcoxon rank-sum
Missing	38	18	56	
NK cell (%s)				0.683
Mean (SD)	1.9 (3.2)	1.9 (3.8)	1.9 (3.5)	Wilcoxon rank-sum
Missing	38	18	56	
Granulocyte (%s)				0.911
Mean (SD)	74.3 (9.7)	74.3 (10.8)	74.3 (10.3)	Wilcoxon rank-sum
Missing	38	18	56	

839

840 **Figures**

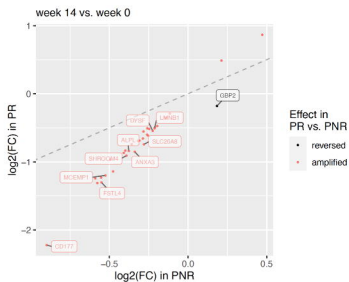
1a



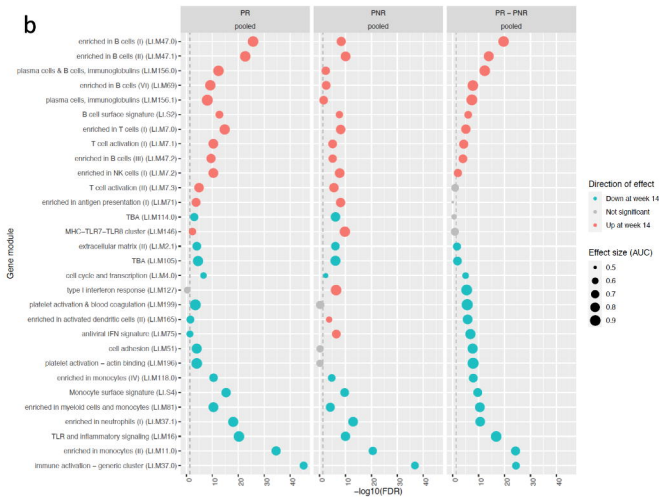
b



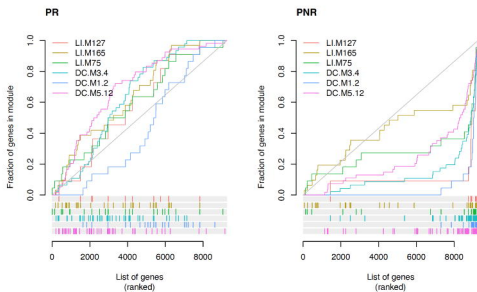
2a



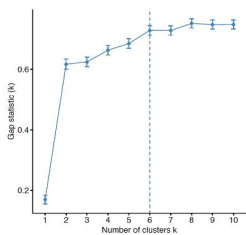
b



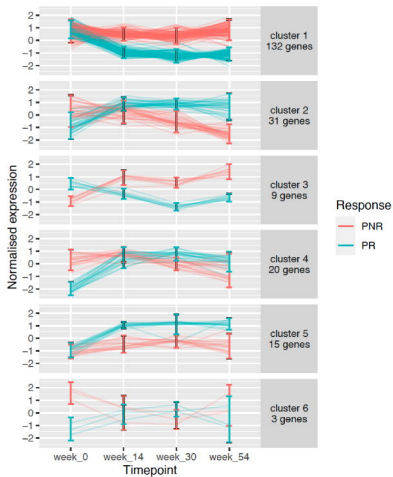
c



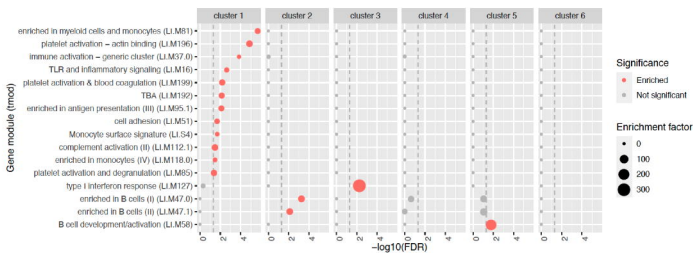
3a



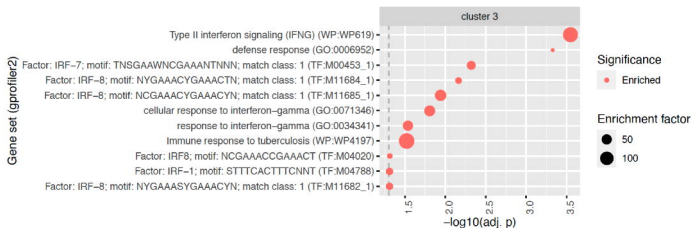
b



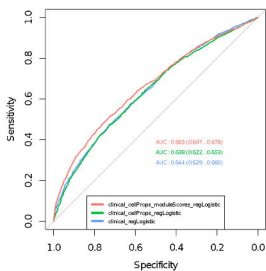
c



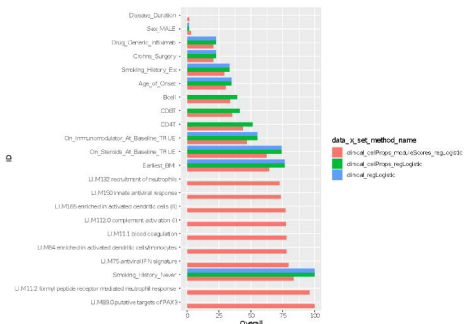
d



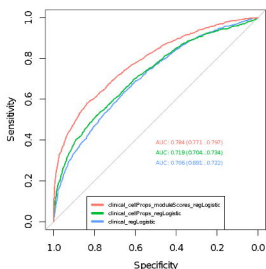
4a



b



c



d

